REPORT DOCUMENTATION PAGE

OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL/01 OCT 92 TO 30 SEP 94

4. TITLE AND SUBTITLE

(HBCU/MI FY92-FY93) A CLUSTER WORKSTATION
FACILITY FOR RESEARCH IN COMPUTER SCIENCE (U)

5. FUNDING NUMBERS

6. AUTHOR(S)

Professor Thomas Mathew

4276/AS
F49620-93-1-0001

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Maryland
Baltimore County
Baltimore MD 21228

8. PERFORMING ORGANIZATION REPORT NUMBER

AFOSR-TR- 95 0078

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NM
110 DUNCAN AVE, SUITE B115
BOLLING AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

F49620-93-1-0001

11. SUPPLEMENTARY NOTES

DTIC
ELECTE
FEB 23 1995
S G D

12a. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

12b. DISTRIBUTION CODE

UL

13. ABSTRACT (Maximum 200 words)

The proposal addressed some statistical inference problems in the areas of multivariate calibration and meta-analysis. Multivariate calibration deals with using the statistical relationship between a response variable and an explanatory variable for statistical inference concerning an unknown value of the explanatory variable using available data. The problems that have been solved deal with the construction of confidence regions for the unknown value of the explanatory variable. Satisfactory solutions to some open problems in this area have been obtained. Meta-analysis deals with combining several independent tests concerning a common parameter. Five papers have been written based on the proposal, of which three are already published. In addition, two papers are currently under preparation.

DTIC QUALITY INSPECTED 4

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR(SAME AS REPORT)

Statistical Inference Problems in Some Multivariate Linear Models With Applications to Multivariate Calibration and Meta-Analysis (AFOSR F49620-93-1-0001)

Thomas Mathew

University of Maryland Baltimore County
Baltimore, Maryland 21228

Air Force Office of Scientific Research
110 Duncan Avenue, Suite B115
Bolling AFB, DC 20332

Accesion For		
NTIS CRA&I		X
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

19950214 007

Distribution Unlimited

Abstract: The proposal addressed some statistical inference problems in the areas of multivariate calibration and meta-analysis. Multivariate calibration deals with using the statistical relationship between a response variable and an explanatory variable for statistical inference concerning an unknown value of the explanatory variable using available data. The problems that have been solved deal with the construction of confidence regions for the unknown value of the explanatory variable. Satisfactory solutions to some open problems in this area have been obtained. Meta-analysis deals with combining information from several independent sources to arrive at an overall combined decision. The problems that have been solved in this context deal with combining several independent tests concerning a common parameter. Five papers have been written based on the proposal, of which three are already published. In addition, two papers are currently under preparation.

calibration, confidence region, meta-analysis
multivariate linear model

4

UNCLASSIFIED   UNCLASSIFIED   UNCLASSIFIED   SAR

The problems formulated in the proposal were in the areas of multivariate calibration and meta-analysis, when the data can be modeled using a multivariate linear model. What follows is a very brief description of the above problems, along with a summary of the work completed and papers prepared under my AFOSR Grant (AFOSR F49620-93-1-0001). I would like to mention at the very beginning that among the work completed under the grant, I consider my work on multivariate calibration to be quite significant. I anticipate that the results I have obtained in this context will be widely used in practical problems dealing with calibration, notably in chemometrics and in spectroscopy. (In response to a request from the Program Director Dr. Jon Sjogren, I selected a paper on calibration – paper #4 below – as my best research accomplishment since October 1993).

Statistical calibration deals with procedures to determine the characteristics of an object (which are difficult to measure directly), using easily obtained measurements on related variables, by means of a statistical model that relates the characteristics of interest with the variables on which measurements can be made. My work in this area, funded by AFOSR, has been on the construction of confidence regions. The area of meta-analysis deals with quantitative procedures for combining information from several independent sources. In this context, the problems formulated in the proposal dealt with combining several independent tests. The following papers have been completed under the grant (part of the work reported in papers #1 and #2 below have been completed under my earlier AFOSR grant, AFOSR 89-0237):

1. (With S. Kasala) An exact confidence region in multivariate calibration. Appeared in *The Annals of Statistics*, 1994.

2. (With L. Zhou) Combining independent tests in multivariate linear models. Appeared in *Journal of Multivariate Analysis*, 1994.

3. (With A. Khuri and D. Nel) A test to determine closeness of multivariate Satterthwaite's approximation. Appeared in *Journal of Multivariate Analysis*, 1994.

4. (With W. Zha) A conservative confidence region in multivariate calibration. Submitted to *The Annals of Statistics*.

5. (With K. Nordström) An inequality for a measure of deviation in linear models. Submitted to *The American Statistician*.

In addition, the following papers are under preparation.

6. (With W. Zha). Multiple use confidence regions in multivariate calibration.

7. (With R. J. Kelly). Nonnegative estimation of variance components in some time

series models.

The research work in papers #1, #4 and #6 deal with the construction of confidence regions in the the statistical calibration problem. The set up is as follows. Consider the $N \times p$ matrix Y of observations following the multivariate linear model

$$Y = BX + E,$$

where B is a $p \times m$ matrix of unknown parameters, X is a known $m \times p$ matrix and E is a matrix of error terms. The rows of E are assumed to be independent having the multivariate normal distribution $N(0, \Sigma)$, the covariance matrix Σ being unknown. The N rows of Y represent N independent observations obtained based on known design points given by the different rows of X. If $\mathbf{y}_0$ represents a future observation corresponding to an unknown design point θ, we get the model

$$\mathbf{y}_0 = B\theta + \mathbf{e}_0,$$

where the error term $\mathbf{e}_0$ is again assumed to follow the $N(0, \Sigma)$ distribution, and is independent of E. Thus the calibration problem consists of estimating B and Σ using Y and using it, along with $\mathbf{y}_0$, to obtain information about θ. We have addressed the problem of constructing confidence regions for θ. In spite of the fact that statistical calibration is a widely used procedure, a satisfactory solution has not been available for this problem, except asymptotically. The procedures that were available in the finite sample case have serious drawbacks - some of the confidence regions could even be empty! We have succeeded in deriving confidence regions that are intuitively appealing, are applicable to finite samples and are free of the drawbacks associated with the existing procedures. Furthermore, when applied to actual data, our confidence regions turned out to be much more satisfactory compared to those based on the existing procedures. The examples and simulation results indicate that we finally have a satisfactory solution to this problem, applicable to small samples.

The problem addressed in paper #2 is that of combining independent tests regarding a common parameter that occurs in the mean of several multivariate linear models. This is a problem in meta-analysis and has important applications in the social and behavioral sciences. This area of research also has applications in industrial robotics, commercial air traffic control and military operations where sensory data from multiple sensors are very often available and the problem is to combine diverse sets of sensed information. When the data can be modelled using a linear model, our results in paper #2 achieve this for hypotheses testing problems. Our results in this regard generalize the corresponding univariate results in the papers by Zhou and Mathew (1993, *Journal of the American Statistical Association*) and Mathew, Sinha and Zhou (1993, *Journal of the American Statistical Association*).

Satterthwaite's approximation of the distribution of a linear combination of independent means squares is a well known procedure in the analysis of univariate balanced mixed models. The multivariate analogue of this approximation is not so well known. This approximation applies to a nonnegative linear combination of independent Wishart matrices. In the article #3, we study the closeness of this approximation. A necessary and sufficient condition for the approximation to be exact is presented. A multivariate test statistic is then developed to detect any significant departure from this condition.

In the article #5, we establish a matrix inequality that provides an upper bound for a quadratic form that involves the difference between two linear unbiased estimators of the same linear parametric function in a general linear model. Various special cases of the inequality are discussed. Certain inequalities that arise in the problem of outlier detection and the problem of combining information from two independent linear models come out as special cases.

The work that is being completed in paper #7 deals with the nonnegative estimation of variance components in some time series models when measurement noise is present. The standard estimator in the time series literature is the maximum likelihood estimator, which can be computationally intense. We cast the time series model in direct correspondence with a mixed effects model and then use the results in the papers by Kelly and Mathew (*Journal of the Royal Statistical Society*, 1993, and *Technometrics*, 1994) in order to arrive at nonnegative estimators of variance components that have explicit analytic expressions, and hence are computationally easy. Furthermore, simulation results show that our estimators exhibit excellent performance (in terms of mean squared error) compared to the maximum likelihood estimator.

As a concluding remark, I would like to add that the most important research accomplishment among the above is the work on calibration, as mentioned at the beginning of this report. The results that I obtained in this context (in papers #1, #4 and #6) are exclusively on the construction of confidence regions. Papers #4 and #6 are jointly with a graduate student, W. Zha, also supported on the AFOSR grant, and will form part of her Ph. D. dissertation. (She will be graduating in the summer of 1995). Along with another graduate student, I am currently working on some point estimation problems that arise in multivariate calibration.